CIS 3362 8/31/22

makers	breakers
substitution	→ freq
null char, code words	→ adjustment

Vigenere Cipher (Unbreakable Code!)

Plaintext : IT IS VERY | HOT OUTSIDE
Keyword : KN IGHTSK | NIGHTSKNIG

	8	19	8	18	21	4	17	24
Shift by +	10	13	8	6	7	19	18	10
	18	32	16	24	28	23	35	34
		↓			↓		↓	↓
		6			2		9	8

18, 6, 16, 24, 2, 23, 9, 8
S G Q Y C X J I ...

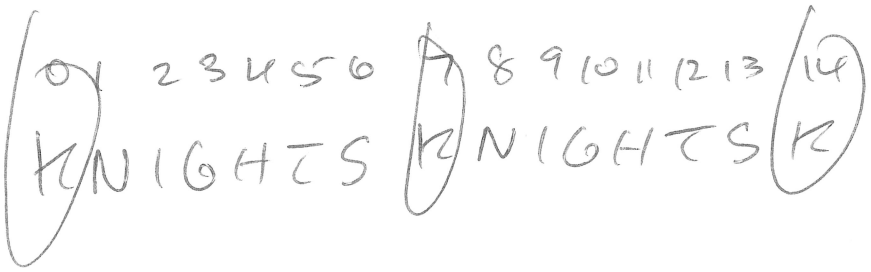
advantages: (1) easier to remember
*(2) freq info distorted
2 T's can encrypt to 2 diff letters
2 diff letters can encrypt to same letter

disadvantage: (1) easier to "guess" at a keyword.

TO DO: Code up Java version encryption

Breaking Vigenere

letter freq not preserved BUT some plaintext chars are shifted by the same amount!



letters pos 0,7,14,... same shift
letters pos 1,8,15,... same shift
⋮
letters pos 6,13,20,... same shift

We have |k| # of bins where each bin has correct letter frequencies.

Issues - bins will have fewer chars than for reg sub.

- won't get digraph, trigraph info!

How to deal w/ issues

(1) repeated digrams trigrams

Imagine I use the same word 5 times.

keyword length was 8.

list 5 random #s from set {0,1,2,3,4,5,6,7} each chosen independently. What's probability they are all different?

6, 3, 4, 1, 7

$$1 \times \frac{7}{8} \times \frac{6}{8} \times \frac{5}{8} \times \frac{4}{8} \cong 70\%$$

80% chance that the repeated word is encrypted identically!!!

Kasiski - repeated ditrigram

⇒ likely means same plaintext

⇒ difference in the indexes is divisible by the keyword length.

(1) find all repeats

(2) record all differences in index

(3) GCD all #s in step.